Problem Statement : For a given dataset (problem) which is the best classification algorithm (as per accuracy)

Pick up a dataset of your choice or the one attached with mail (Preferable)

Ask any two questions on the dataset of your choice and provide answers for the same.

For instance for the given dataset questions can be as follows.

Q1) What are the most common emotions/words used by Males and Females?

Q2) Which gender makes more typos in their tweets?

Take up three classification algorithms of your own choice and build three respective Machine learning models. Compare the Accuracy of all three and suggest which ML algorithms suits best for the given problem.

NOTE  : For the given dataset "Gender" will be your dependent variable.

**Evaluation will be done on following points:**

1)  Exploratory data analysis (with visualization) and Data Cleaning if required

2) Questions asked on dataset and answers for the same with brief explanation

3) Feature Selection and feature Engineering if required depending on the dataset

4) Ensemble Machine learning Modelling (3 Classification Algorithms)

5) Accuracy calculation

6) Summarised write up at the end

**OPTIONAL REQUIREMENT :** It will be appreciated if any one algorithm is built from scratch instead of using a library.

Please explain all your steps with clear details and comments. Do mention which are your Independent and dependent variables on the dataset

**Prepare a PDF/Word Document at the end with Summary of this project and submit.**

The last date of submission is **9th Jan 2021.**

**Submission:**

1. Please submit your projects individually to [event@verzeo.in](mailto:event@verzeo.in) and in your respective classrooms.
2. Any project submitted to [support@verzeo.in](mailto:support@verzeo.in) will not be evaluated.
3. The subject line for your submission must be in the following format- **ML-MAJOR-NOV-** **GOOGLE CLASSROOM NAME**.
4. Projects without a proper subject line will not be evaluated.